

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-105766

(43)Date of publication of application : 11.04.2000

(51)Int.Cl.

G06F 17/30

(21)Application number : 10-274539

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 29.09.1998

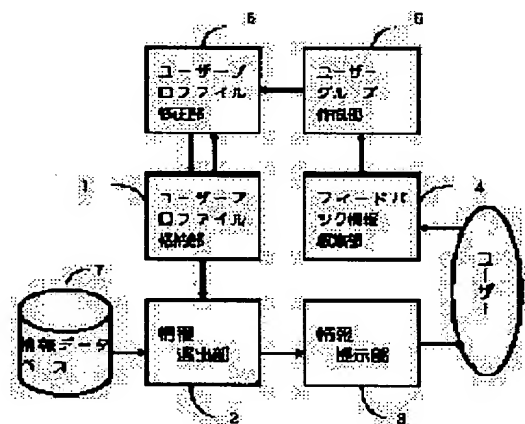
(72)Inventor : SAKAI TETSUYA  
UEHARA TATSUYA  
SUMITA KAZUO

## (54) DEVICE AND METHOD FOR INFORMATION FILTERING, AND STORAGE MEDIUM

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To allow the number of users to be increased and to obtain a relevance feedback which is effective to the individual users by generating user groups according to gathered feedback information and correcting user profiles for every user group according to the gathered feedback information.

**SOLUTION:** An information selection part 2 select information which seems to be proper out of an information data base 7 according to an initial user profile stored in a user profile storage part 1 and passes it to an information presentation part 3. A user group generation part 5 generates user groups wherein users who are similar in request and interest are gathered according to feedback information from the respective users. A user profile correction part 6 provides relevance feedback by the user groups and stores corrected user profiles in the user profile storage part 1.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開 2000-105766  
(P2000-105766A)  
(43) 公開日 平成12年4月11日 (2000. 4. 11)

(51) Int. Cl. 7

識別記号

F I

テーマコード (参考)

G 0 6 F 17/30

G 0 6 F 15/403 3 4 0 A 5B075

審査請求 未請求 請求項の数 4

〇 L

(全 1 1 頁)

(21) 出願番号 特願平10-274539

(22) 出願日 平成10年9月29日 (1998. 9. 29)

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 酒井 哲也

神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内

(72) 発明者 上原 龍也

神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内

(74) 代理人 100083161

弁理士 外川 英明

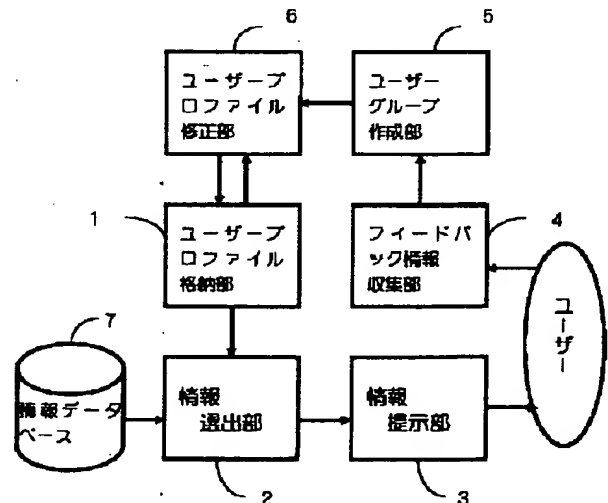
最終頁に続く

(54) 【発明の名称】 情報フィルタリング装置および方法および記憶媒体

(57) 【要約】

【課題】 ユーザー数が増大した場合、検索結果に対するフィードバックを各ユーザーから収集し、それを各ユーザーのユーザープロフィールに反映させることは困難であった。

【解決手段】 ユーザー数が増大しても、その中に共通の要求・興味をもったユーザーが存在する場合に、ユーザー毎ではなくユーザーグループ毎にrelevance feedbackを行うことにより、フィードバックの処理量が軽減され、かつ個々のユーザーから見ても効果的なフィードバックを行うことができる。また、複数のグループ化方法により様々なユーザーグループを定義し、これらに対してrelevance feedbackを行った場合のフィルタリング精度を実際にテストしてみることで、relevance feedbackの観点から効果的なユーザーグループを構成することができる。



## 【特許請求の範囲】

【請求項 1】 ユーザーの要求・興味を表すユーザープロフィールを格納するプロフィール格納手段と、  
このプロフィール格納手段に格納されたユーザープロフィールに基づいてユーザーに配信すべき情報を、データベースから選出する選出手段と、  
この選出手段によって選出された情報に対する個々のユーザーのフィードバック情報を収集するための収集手段と、

この収集手段によって収集されたフィードバック情報に基づいて、要求・興味の類似するユーザーを集めたユーザーグループを作成する作成手段と、  
この作成手段で作成されたユーザーグループ毎のユーザープロフィールを、前記収集手段によって収集されたフィードバック情報をもとに修正するプロフィール修正手段とを具備することを特徴とする情報フィルタリング装置。

【請求項 2】 前記プロフィール修正手段によって修正されたユーザープロフィールの精度を評価する評価手段をさらに有し、

前記作成手段は、複数のユーザーグループ作成方法で、複数のユーザーグループを作成し、前記評価手段によって評価された最も精度の高いユーザープロフィールを持つユーザーグループのユーザーグループ作成方法を採用することにより最終的なユーザーグループを決定することを特徴とする請求項 1 記載の情報フィルタリング装置。

【請求項 3】 格納されたユーザーの要求・興味を表すユーザープロフィールに基づいてユーザーに配信すべき情報を、データベースから選出し、  
この選出された情報に対する個々のユーザーのフィードバック情報を収集し、

この収集されたフィードバック情報に基づいて、要求・興味の類似するユーザーを集めたユーザーグループを作成し、

この作成されたユーザーグループ毎のユーザープロフィールを、収集された前記フィードバック情報をもとに修正することを特徴とする情報フィルタリング方法。

【請求項 4】 ユーザーの要求・興味を満たす情報を選出して配信するプログラムをコンピュータに読み取り可能なように記憶させた記憶媒体であって、

格納されたユーザーの要求・興味を表すユーザープロフィールに基づいてユーザーに配信すべき情報を、データベースから選出し、

この選出された情報に対する個々のユーザーのフィードバック情報を収集し、

この収集されたフィードバック情報に基づいて、要求・興味の類似するユーザーを集めたユーザーグループを作成し、

この作成されたユーザーグループ毎のユーザープロフィールを、収集された前記フィードバック情報をもとに修

正するプログラムを記憶した記憶媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、ユーザーの要求・興味にあった情報を選出して配信する情報フィルタリング装置および方法および記憶媒体に関する。

## 【0002】

【従来の技術】 近年、WWW(World Wide Web)や電子図書館の発展に見られるように、テキスト、音声、画像などの大量の情報に個人がアクセスすることが可能になりつつある。アクセス可能な情報が増えるほど、その中から個人の要求・興味にあった情報を効率よくかつ精度よく選出して提示してくれる情報フィルタリングの技術が重要となる。

【0003】 情報フィルタリングにおいては、個々のユーザーの要求・興味はユーザープロフィールにより表現される。例えばテキスト情報のフィルタリングの場合には、ユーザープロフィールにユーザーの要求・興味に関連する語彙を用いた検索条件を記述する方法などが用いられる。

【0004】 しかし、ユーザーの要求・興味を始めからユーザープロフィールにより精緻に表現することは困難であり、さらに、ユーザーの要求・興味自体、時間や状況とともに推移していくという問題がある。

【0005】 このため情報フィルタリングにおいては、情報検索の分野で古くから研究されている relevance feedback という技術を用いてユーザープロフィールを修正していくことが重要であると考えられる。Relevance feedback は、まず初期プロフィールにより得られたフィルタリング結果に対してユーザーが評価を行い、この評価情報を利用してプロフィールを修正することにより、次のフィルタリングの精度を高めようとするものである。

【0006】 例えばテキスト情報のフィルタリングの場合には、ユーザーにフィルタリング結果の中から「興味のある文書」、「興味のない文書」などをいくつか選択してもらい、これらの本文中から新しい語彙を抽出して、プロフィールに反映させるといった方法が考えられる。テキスト情報を対象とした relevance feedback の具体的手法は、例えば文献（"Introduction to Modern Information Retrieval, Salton & McGill, McGraw-Hill Book Company, 1983）に開示されている。

【0007】 従来の情報フィルタリングシステムでは、複数のユーザーに対してサービスを行う場合、以下のふたつの方法のいずれかをとっていた。第一は、個々のユーザーの興味に応じて個々のユーザープロフィールを修正する方法である。

【0008】 第二は、個々のユーザーに対してユーザープロフィールをもつのではなく、あらかじめいくつかのトピックをシステム側で用意し、ユーザーにこの中から

選択してもらう方法である。

【0009】前者では、ユーザ数の増大に比例してrelevance feedback処理のシステム負荷が増大してしまうという問題がある。後者では、この問題を回避するためにユーザー毎ではなくトピック毎にプロフィールをもつものであり、一般にひとつのトピックプロフィールを複数ユーザーが共有することになる。

【0010】また、複数ユーザーを扱い、ユーザーの要求・興味に合った情報を提供する他の試みとしては、協調フィルタリング (collaborative filtering, social filtering) という技術が知られている。これは、例えば文献 ("Social Information Filtering: Algorithms for Automating Word of Mouth, Upendra Shardanand and Pattie Maes, ACM CHI'95 Proceedings, 1995) に開示されているように、ユーザーのもつ興味の間の類似度をもとに、あるユーザーがおもしろいと判断した情報は、そのユーザーと似たユーザーにも提供しようとするものである。

【0011】

【発明が解決しようとする課題】しかし、同一トピックを選択した複数ユーザーの要求・興味は必ずしも似ているとは限らないため、これらのユーザーから得たフィードバック情報を利用して共通のトピックプロフィールを修正することは難しい。

【0012】例えば、ユーザーAとユーザーBが共に「DVD」というトピックを選択し、ユーザーAのフィードバック情報からユーザーAは、「DVDプレーヤー」などハード面に興味があるが、ユーザーBのフィードバック情報からユーザーBは、「映画」「DVDタイトル」などソフト面に興味があることがわかる場合、両者を満足させるように「DVD」のプロフィールを修正することは難しい。

【0013】本発明は、このような状況に鑑み、ユーザー数の増大を許容し、かつ個々のユーザーにとって効果的なrelevance feedbackを実現することを目的としたものである。

【0014】また、協調フィルタリングの方法では、どのユーザーも評価を行ったことのない未知の情報は扱うことができない。これに対しRelevance feedbackは、既知の情報に対するユーザーの評価情報を利用してプロフィールの修正を行い、これにより未知の情報に対するフィルタリングの精度を向上することを目的とする。

【0015】

【課題を解決するための手段】これらの課題を解決するために、本願発明はユーザーの要求・興味を表すユーザープロフィールを格納するプロフィール格納手段と、このプロフィール格納手段に格納されたユーザープロフィールに基づいてユーザーに配信すべき情報を、データベースから選出する選出手段と、この選出手段によって選出された情報に対する個々のユーザーのフィードバック

情報を収集するための収集手段と、この収集手段によって収集されたフィードバック情報に基づいて、要求・興味の類似するユーザーを集めたユーザーグループを作成する作成手段と、この作成手段で作成されたユーザーグループ毎のユーザープロフィールを、前記収集手段によって収集されたフィードバック情報をもとに修正するプロフィール修正手段とを具備することを特徴とする。

【0016】このような本願発明により、ユーザー数が増大しても、その中に共通の要求・興味をもったユーザーが存在する場合に、ユーザー毎ではなくユーザーグループ毎にrelevance feedbackを行うことにより、フィードバックの処理量が軽減され、かつ個々のユーザーから見ても効果的なフィードバックを行うことができる。

【0017】また、複数のグループ化方法により様々なユーザーグループを定義し、これらに対してrelevance feedbackを行った場合のフィルタリング精度を実際にテストしてみることで、relevance feedbackの観点から効果的なユーザーグループを構成することができる。

【0018】

【発明の実施の形態】(実施例1) 図1に、本実施例における情報フィルタリング装置の構成を示す。矢印はデータの流れを表している。本装置は、ユーザープロフィール格納部1、情報選出部2、情報提示部3、フィードバック情報収集部4、ユーザーグループ作成部5、ユーザープロフィール修正部6により構成される。

【0019】ここで、ユーザープロフィール格納部1は主記憶や外部記憶装置などに、情報提示部3はユーザーに情報を配信するための計算機ネットワークや表示するための計算機ディスプレイなどの出力装置に、フィードバック情報収集部4はユーザーからのフィルタリング結果に対する評価情報の入力を受けつける入力装置に、その他の構成要素は計算機のプログラムに対応する。

【0020】以下に、図1に沿って本発明における全体的な処理の流れを説明する。はじめに情報選出部2は、ユーザープロフィール格納部1に格納されている初期ユーザープロフィールをもとに、情報フィルタリングの対象となる情報データベース7から個々のユーザーにふさわしいと思われる情報を選出し、これを情報提示部3にわたす。各ユーザーは、自分に提示された初期フィルタリング結果を見て、これに対する評価情報をフィードバック情報収集部4をとおして入力する。

【0021】次に、ユーザーグループ作成部5は、各ユーザーのフィードバック情報をもとに、要求・興味が類似していると思われるユーザーを集めたユーザーグループを作成する。最後に、ユーザープロフィール修正部6は、ユーザーグループ毎にrelevance feedbackを行い、修正されたユーザープロフィールをユーザープロフィール格納部1に格納する。

【0022】本実施例の従来の情報フィルタリングシステムとの相違点は、フィードバック情報収集部4とユー

ザープロファイル修正部 6 の間にユーザーグループ作成部 5 を有する点である。

【0023】図 2 に、本実施例における情報選出部 2 の処理の流れの一例を示す。情報選出部 2 は、各ユーザーごとにユーザープロファイルをユーザープロファイル格納部 1 から取り出し (S21)、ユーザープロファイルと情報との類似度を算出し (S22)、類似度の高いもののみを情報提示部 3 にわたすことにより情報を絞り込む (S23)。

【0024】ここで、本発明は類似度算出の具体的手法を問うものではなく、例えばテキスト情報のフィルタリングの場合には、ユーザープロファイルとテキスト情報の双方を単語の重みつきベクトルで表現し、このベクトル間の余弦により類似度を決定する文献 ("Introduction to Modern Information Retrieval, Salton & McGill, McGraw-Hill Book Company, 1983) などに開示されている手法などを用いればよい。

【0025】例えば、ユーザーが「DVD」に関する話題について興味がある場合には、このユーザーの初期ユーザープロファイルを図 3 (a) のような、「DV」、「D」、「デジタル」、「ディスク」という単語の集合で表しておき、これらの語が出現する頻度に応じて各テキスト情報の類似度を算出することが考えられる。

【0026】図 4 に、本実施例におけるフィードバック情報収集部 4 の処理の流れの一例を示す。フィードバック情報収集部 4 は、各ユーザーに、初期フィルタリング結果に対する評価情報を入力させ (S41)、これをユーザーグループ作成部 5 に渡す (S42)。

【0027】図 5 に、各ユーザーのフィードバック情報の一例を示す。この例は、簡単のためにユーザー数を 4 人としており、各ユーザーが情報 1～5 に対して 2 段階評価を行った様子を表している。例えばユーザー 1 は、情報 2 および情報 3 に対しては「興味がある」という判定を行ったが、情報 1、情報 4 および情報 5 に対しては「興味がない」という判定を行ったことを示している。

【0028】あるいは、「興味がある」という情報だけを収集して、「興味がある」という評価がされなかった残りの情報に対しては「0」とするという方法も考えられる。さらに、以上の変形例として、各情報に対する評価を多段階評価にすることも考えられる。本発明は、以上で説明したようなフィードバック情報の形態を問うものではない。

【0029】図 6 に、本実施例におけるユーザーグループ作成部 5 の処理の流れの一例を示す。ユーザーグループ作成部 5 は、図 5 で例示したような各ユーザーのフィードバック情報をフィードバック情報収集部 4 から受け取り (S61)、これをもとに、要求・興味が類似していると思われるユーザーを集めてユーザーグループを作成し (S62)、このユーザーグループ情報をユーザープロファイル修正部 6 にわたす (S63)。

【0030】本発明は、ユーザーグループのグループ化方法を問うものではないが、図 5 のようなフィードバック情報が与えられた場合のグループ化の方法の簡単な例を図 7 を利用して説明する。図 7 では、図 5 で示したフィードバック情報に対してクラスター分析を行うことによりユーザーグループを作成している。

【0031】まず図 5 において、各ユーザーベクトル間の類似度を Jaccard 係数 (対応する要素が 1-1 である個数を  $a$ 、1-0 である個数を  $b$ 、0-1 である個数を  $c$ 、0-0 である個数を  $d$  としたとき、 $a / (a + b + c)$  で算出される) という尺度により算出する。

【0032】例えばユーザ 1 とユーザ 2 の類似度は、 $a = 1$ 、 $b = 1$ 、 $c = 2$  より  $\text{sim}(1, 2) = \text{sim}(2, 1) = 1 / 4$  となる。同様に、 $\text{sim}(3, 1) = 0$ 、 $\text{sim}(3, 2) = 1 / 4$ 、 $\text{sim}(4, 1) = 1 / 4$ 、 $\text{sim}(4, 2) = 1 / 2$ 、 $\text{sim}(4, 3) = 1 / 4$  となる。

【0033】以上のうち、類似度が最も高かったユーザ対を同一グループとしてグループ化する。これにより、ユーザ 2 とユーザ 4 を同一グループとし、ユーザーグループは、(2, 4) (1) (3) の 3 つになる。

【0034】次に、例えば UPGMA 法 (算術平均を用いた対グループ法) という手法を用いて、 $\text{sim}((2, 4), 1) = (\text{sim}(2, 1) + \text{sim}(4, 1)) / 2 = 1 / 4$  のように、ユーザーグループ間の類似度を計算する。その他の組合せについては図 7 に示したようになる。

【0035】ここで、 $\text{sim}((2, 4), 1) = 1 / 4$ 、 $\text{sim}((2, 4), 3) = 1 / 4$ 、 $\text{sim}(1, 3) = 0$  のうち類似度の大きいのは  $\text{sim}((2, 4), 1)$  あるいは  $\text{sim}((2, 4), 3)$  である。仮に後者を選べば、新しいユーザーグループ (2, 3, 4) ができることになる。ちなみに、 $\text{sim}((2, 3, 4), 1) = 1 / 6$  となる。

【0036】以上のようにして、(1) (2) (3) (4)  $\rightarrow$  (2, 4) (1) (3)  $\rightarrow$  (2, 3, 4) (1)  $\rightarrow$  (1, 2, 3, 4) のようにクラスター分析を行うことができるので、これから指定したクラス数あるいは指定した類似度のもとでのユーザーグループ群を決定することができる。

【0037】以上では、グループ化の手法の一例としてクラスター分析について説明したが、このようなグループ化の手法自体は例えば文献 ("Cluster Analysis for Researchers, Romesburg, Robert E. Krieger Publishing Company, Malabar, Florida, 1989) に開示された公知の技術であり、本発明の主眼ではない。例えば、Jaccard 係数の他にも Sorenson 係数というものがあり、また、UPGMA 法の他にも Ward 法などがある。

【0038】図 8 に、本実施例におけるユーザープロファイル修正部 6 の処理の流れの一例を示す。本実施例におけるユーザープロファイル修正部 6 は、従来の情報フィルタリングシステムのようにユーザー毎に relevance feedback を行うのではなく、ユーザーグループ毎に relevance feedback を行うことを特徴とする。具体的には、ユーザープロファイル修正部は、各ユーザーグループに

7  
対して以下の処理を行う。

【0039】まずユーザーグループ情報をユーザーグループ作成部5から受け取り(S81)、ユーザーグループに属する各ユーザーのフィードバック情報を用いて、「ユーザーグループのフィードバック情報」を決定する(S82)。例えば、図5および図7の例の場合、ユーザーグループ(2,4)に対しては、ユーザー2のベクトル(01011)とユーザー4のベクトル(00111)をビット毎にORをとることにより(01111)とするなどして、ひとりのユーザーのフィードバック情報と同様の形態のフィードバック情報を作成する。

【0040】他には、例えば多数決で、個々の情報を「興味がある」と見做すか否かを決定する方法が考えられる。ただし本発明は、以上で例をあげて説明したような、複数ユーザーのフィードバック情報をひとつに統合する具体的な手法を問うものではなく、これは本発明の主眼ではない。これにより、従来のrelevance feedback手法は本発明にそのまま適用可能となる。

【0041】ユーザーグループに対するフィードバック情報が決定されると、ユーザープロファイル修正部6は、そのユーザーグループに属する各ユーザーのユーザープロファイルに対して、上記フィードバック情報を反映させる(S83)。本発明は、具体的なrelevance feedback手法を問うものではないが、テキスト情報のフィルタリングなどの場合には、フィードバック情報の反映とは、ユーザープロファイルに記述された単語の重みの調整や、新しい単語の追加などを意味する。

【0042】例えば、図3(a)のような初期ユーザープロファイルに対して、フィードバック情報をもとに新しい単語を追加した例が図3(b)である。本発明では、このようなユーザープロファイルの修正処理を、複数のユーザープロファイルに対して一括して行う。

【0043】以上、主にテキスト情報を対象としたフィルタリングにおいて、単語の羅列によりユーザープロファイルを表現した場合を例に説明したが、本発明はフィルタリングの対象となる情報の種類や、ユーザープロファイルの表現形式を限定するものではない。

【0044】(実施例2)図9に、本実施例における情報フィルタリングシステムの構成を示す。本実施例と実施例1との構成上の違いは、ユーザープロファイル評価部8を有する点である。以下、実施例1と異なる点のみについて説明する。

【0045】全体の流れについては、情報選出部2、情報提示部3、フィードバック情報収集部4までは実施例1と同じであるが、それ以降、複数のグループ化手法が用いられる点異なる。すなわち、本実施例では、ユーザーグループ作成部5が、複数のグループ化手法を用いることにより、ユーザーグループ群を複数セット作成して以降の処理にわたす。

【0046】例えば、図7の例において、実施例1の場

合は、例えばユーザーグループ数を2あるいは類似度を $1/4$ と決めることにより、ひとつのユーザーグループ群(2,3,4)(1)が決定されるが、本実施例では、例えばユーザーグループ数を2~3、あるいは類似度を $1/2 \sim 1/4$ と変化させることにより、ふたつのユーザーグループ群(2,4)(1)(3)および(2,3,4)(1)を得ることができる。

【0047】また、図7の説明では、ユーザーベクトル間の類似度の尺度としてJaccard係数を用い、クラスタ分析の手法としてはUPGMA法を用いた場合について説明したが、これら以外のクラスタ分析手法を用いて異なるユーザーグループ群を得ることもできる。さらに、ユーザーを分類するためにクラスタ分析以外の手法を用いてもよい。本実施例では、以上のようなグループ化における様々なバリエーションを総称してグループ化手法と呼ぶ。

【0048】図10に、本実施例におけるユーザープロファイル評価部8の処理の流れの一例を示す。ユーザープロファイル評価部8は、各グループ化手法を以下の手順で評価し、最終的にどのグループ化手法がよいか、ひいてはどのようなユーザーグループ群が適切であるかを決定する。

【0049】ユーザープロファイル評価部8は、あるグループ化手法をもとに修正されたユーザープロファイルの各々をユーザープロファイル格納部1から取り出し(S101)、そのフィルタリング精度を評価する(S102)。ここで、フィルタリング精度の尺度としては、例えば適合率(フィルタリング結果中の何パーセントが実際にユーザーが求めているものであるか)などが考えられる。フィルタリング精度の評価実験には、ユーザーが既に評価を行った情報を含むデータをフィルタリング対象として用いることが考えられる。

【0050】例えば、図5の例のユーザー1は、既に情報1~情報5に対しては興味があるかないかの判定を行っているので、これらを評価用のデータに含めて、情報選出部2による仮のフィルタリング結果を作成し、この上位にユーザー1が興味があると判定した情報2や情報3がくるかを確認することができる。

【0051】ユーザープロファイル評価部8は、個々のユーザープロファイルのフィルタリング精度を算出すると、次にそのグループ化手法の平均的フィルタリング精度を算出する(S103)。これは例えば前述の各ユーザープロファイルの適合率を平均するなどして算出すればよい。

【0052】以上の処理を全てのグループ化手法に対して行い、最後に全グループ化手法のフィルタリング精度を比較し、もっとも効果的と思われるグループ化手法を採用し、これにより最終的なユーザークラスタを定める(S104)。

【0053】以上の流れを、図11の具体例を用いて説

明する。この例では、図5およびユーザー1～4の初期ユーザープロフィールがそれぞれ図の上部に示されている。これらの4人のユーザーに対して、例えば図7のようなクラスター分析を行い、ユーザーグループが3の場合および2の場合を採用することにより、図11の2つの矢印で示されるように、ユーザーグループ群(2, 4)(1)(3)および(2, 3, 4)(1)をもとにした2とおりのグループ化手法によるrelevance feedbackを行うことができる。

【0054】これにより得られた修正後のユーザープロフィールの概念的に表したものが図11の中央に示されている。各長方形の上半分は、初期ユーザープロフィールの時点で既にあった情報を表し、下半分は、relevance feedbackにより新たに付加された情報の部分を表している。

【0055】ここで、修正されたユーザープロフィールのうち左側のユーザー2、ユーザー4の下半分を同じ模様で図示しているのは、グループ化手法1ではユーザー1とユーザー2が同じユーザーグループに属するとみなされ、共通のrelevance feedbackが行われたことを表している。右側のユーザー3、4の下半分の模様が共通であるのも同様のことを表している。

【0056】この例では、グループ化手法1の平均的フィルタリング精度は、図の左側の4つのユーザープロフィールのフィルタリング精度の平均により算出され、グループ化手法2の平均的フィルタリング精度は右側の4つの平均により算出される。そして、例えば前者が後者よりも大きい場合、グループ化手法1のほうが最終的に採用され、ユーザーグループ群は(2, 4)(1)(3)と決定される。

【0057】ユーザーをグループ化する場合、どのような大きさのユーザーグループをいくつくらい作成すればよいのかを予め定めることが難しいが、本手法によれば、個々のユーザーにとってrelevance feedbackの効果が大きくなるように適切なグループ化を行うことができると考えられる。

【0058】(記憶媒体の実施形態)本願発明は、専用のハードウェアを用いずとも、汎用のプロセッサを用いたソフトウェアによる処理で実現することができる。例えば、図2、図4、図6、図8、図10の処理はコンピュータプログラムを用いて実現でき、これをフロッピーディスクやCD-ROMなどの記録媒体を通じてコンピュータに導入して実行させることにより、本願発明を実施することができる。

【0059】

【発明の効果】このような本願発明により、ユーザー数

が増大しても、その中に共通の要求・興味をもったユーザーが存在する場合に、ユーザー毎ではなくユーザーグループ毎にrelevance feedbackを行うことにより、フィードバックの処理量が軽減され、かつ個々のユーザーから見ても効果的なフィードバックを行うことができる。

【0060】また、複数のグループ化方法により様々なユーザーグループを定義し、これらに対してrelevance feedbackを行った場合のフィルタリング精度を実際にテストしてみることにより、relevance feedbackの観点から効果的なユーザーグループを構成することができる。

【図面の簡単な説明】

【図1】本願発明の一実施形態である実施例1の構成を示すブロック図。

【図2】本願発明の一実施形態である実施例1における情報選出部の処理の流れの一例を示す図。

【図3】本願発明の一実施形態である実施例1におけるユーザープロフィールの表現例を示す図。

【図4】本願発明の一実施形態である実施例1におけるフィードバック情報収集部の処理の流れの一例を示す図。

【図5】本願発明の一実施形態である実施例1における各ユーザーのフィードバック情報の一例を示す図。

【図6】本願発明の一実施形態である実施例1におけるユーザーグループ作成部の処理の流れの一例を示す図。

【図7】本願発明の一実施形態である実施例1におけるユーザーグループ作成方法の一例を示す図。

【図8】本願発明の一実施形態である実施例1におけるユーザープロフィール修正部の処理の流れの一例を示す図。

【図9】本願発明の一実施形態である実施例2の構成を示すブロック図。

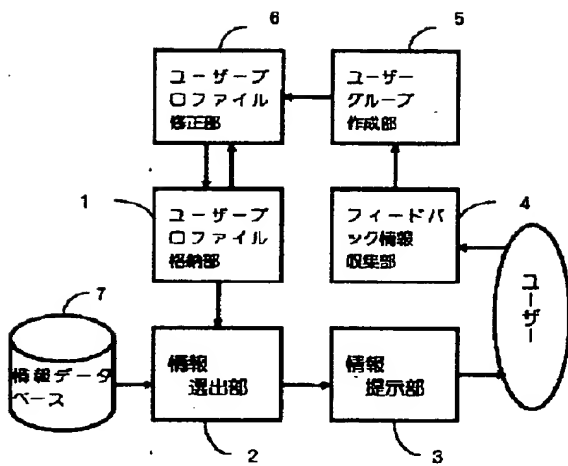
【図10】本願発明の一実施形態である実施例2におけるユーザープロフィール評価部の処理の流れの一例を示す図。

【図11】本願発明の一実施形態である実施例2におけるグループ化手法の比較の流れの概念の一例を示す図。

【符号の説明】

- 1…ユーザープロフィール格納部
- 2…情報選出部
- 3…情報提示部
- 4…フィードバック情報収集部
- 5…ユーザーグループ作成部
- 6…ユーザープロフィール修正部
- 7…情報データベース
- 8…ユーザープロフィール評価部

【図 1】



【図 3】

(a)

DVD, デジタル, ディスク

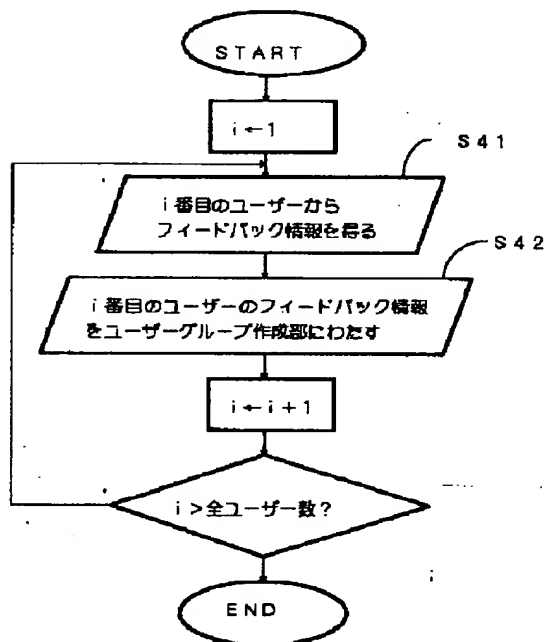
(b)

DVD, デジタル, ディスク  
プレーヤー, ポータブル, DVD-ROM

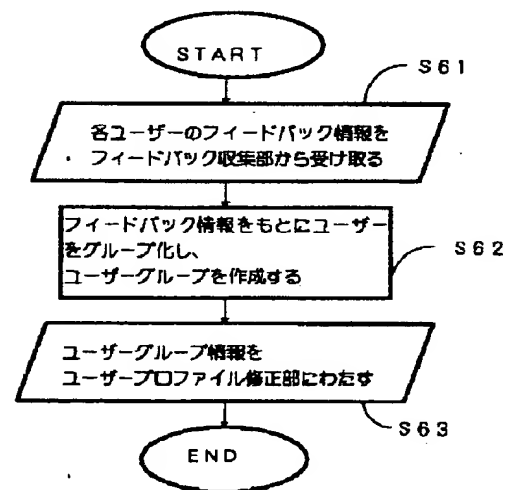
【図 5】

	情報 1	情報 2	情報 3	情報 4	情報 5
ユーザー 1	0	1	1	0	0
ユーザー 2	0	1	0	1	1
ユーザー 3	1	0	0	0	1
ユーザー 4	0	0	1	1	1

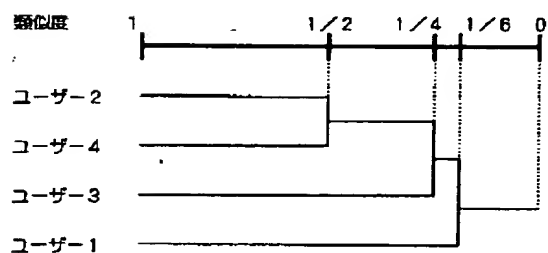
【図 4】



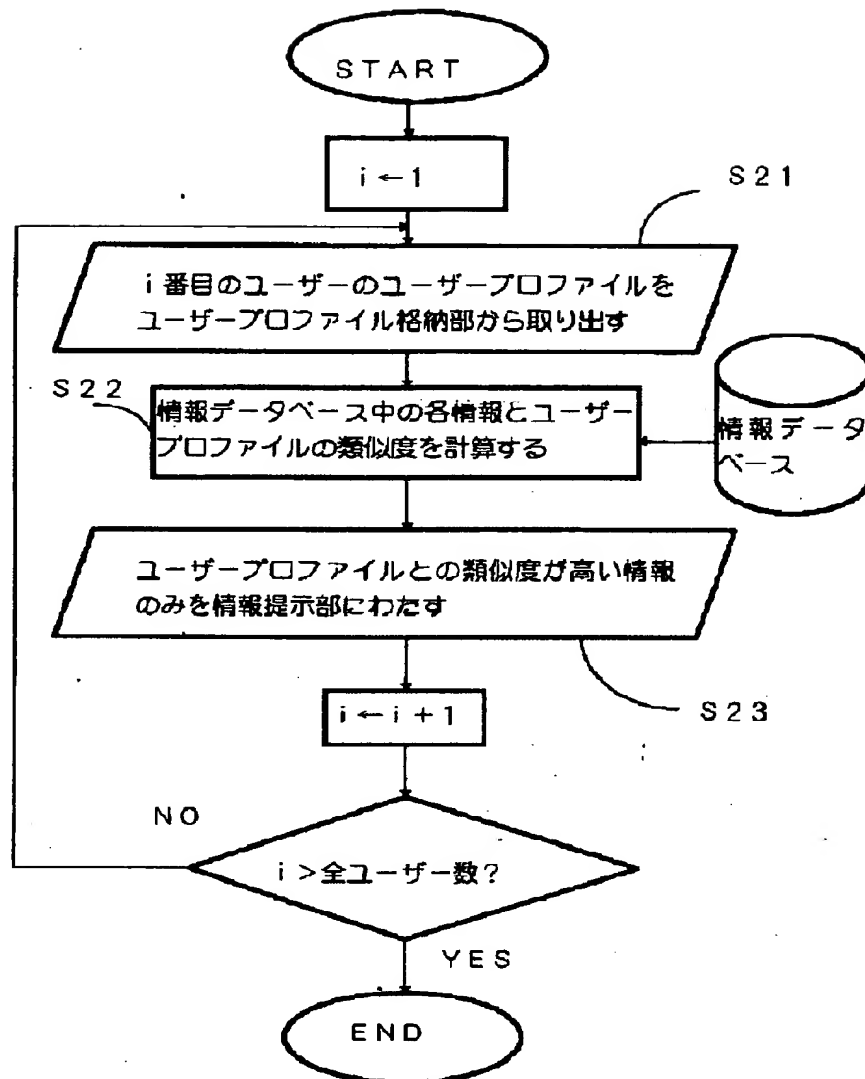
【図 6】



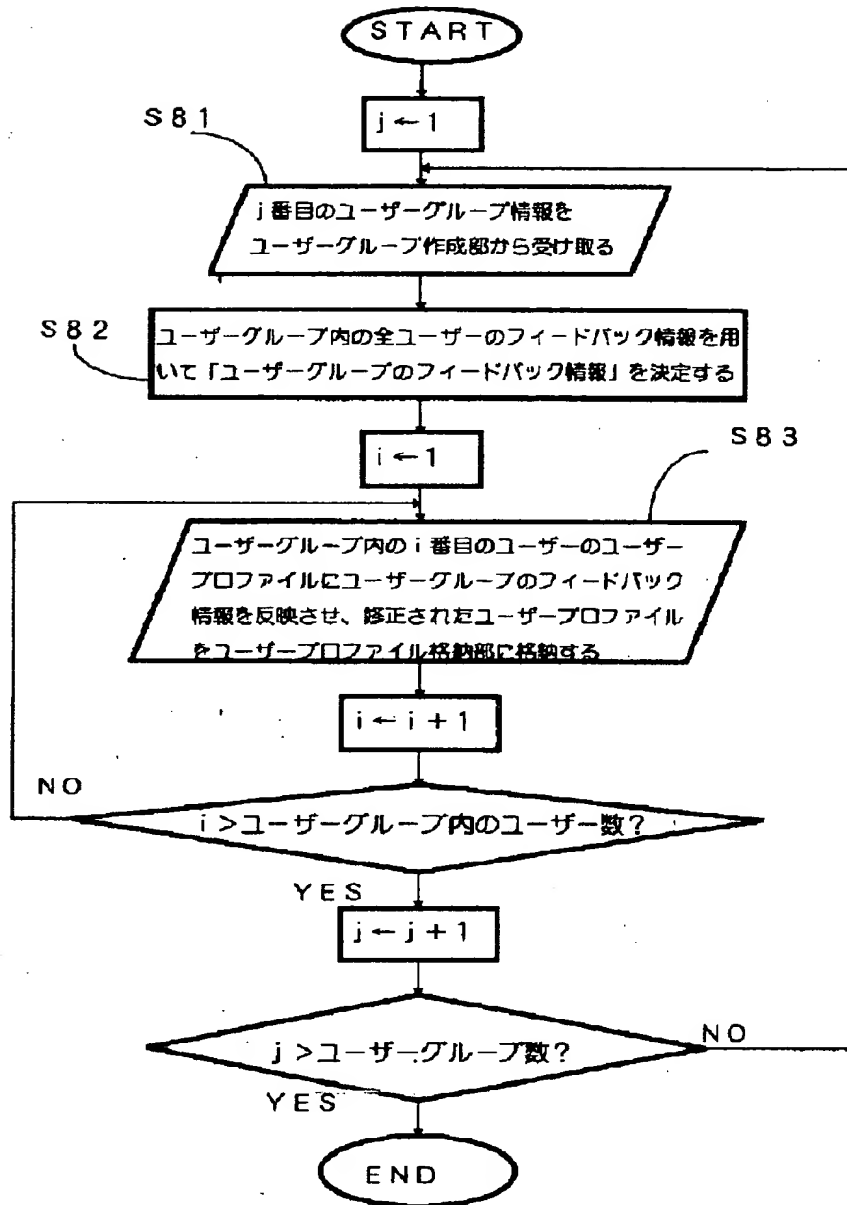
【図 7】



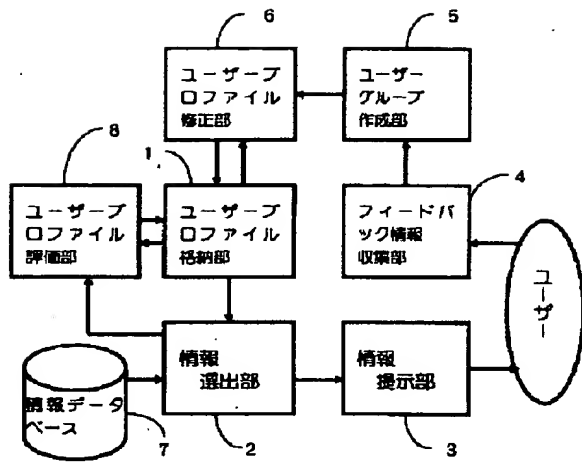
【図2】



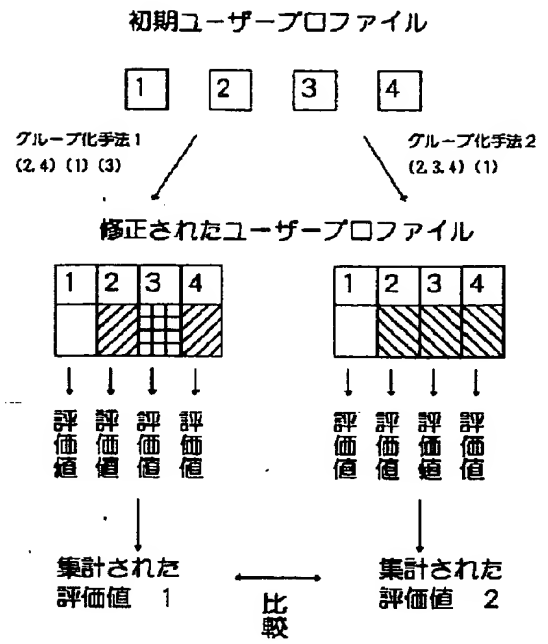
【図 8】



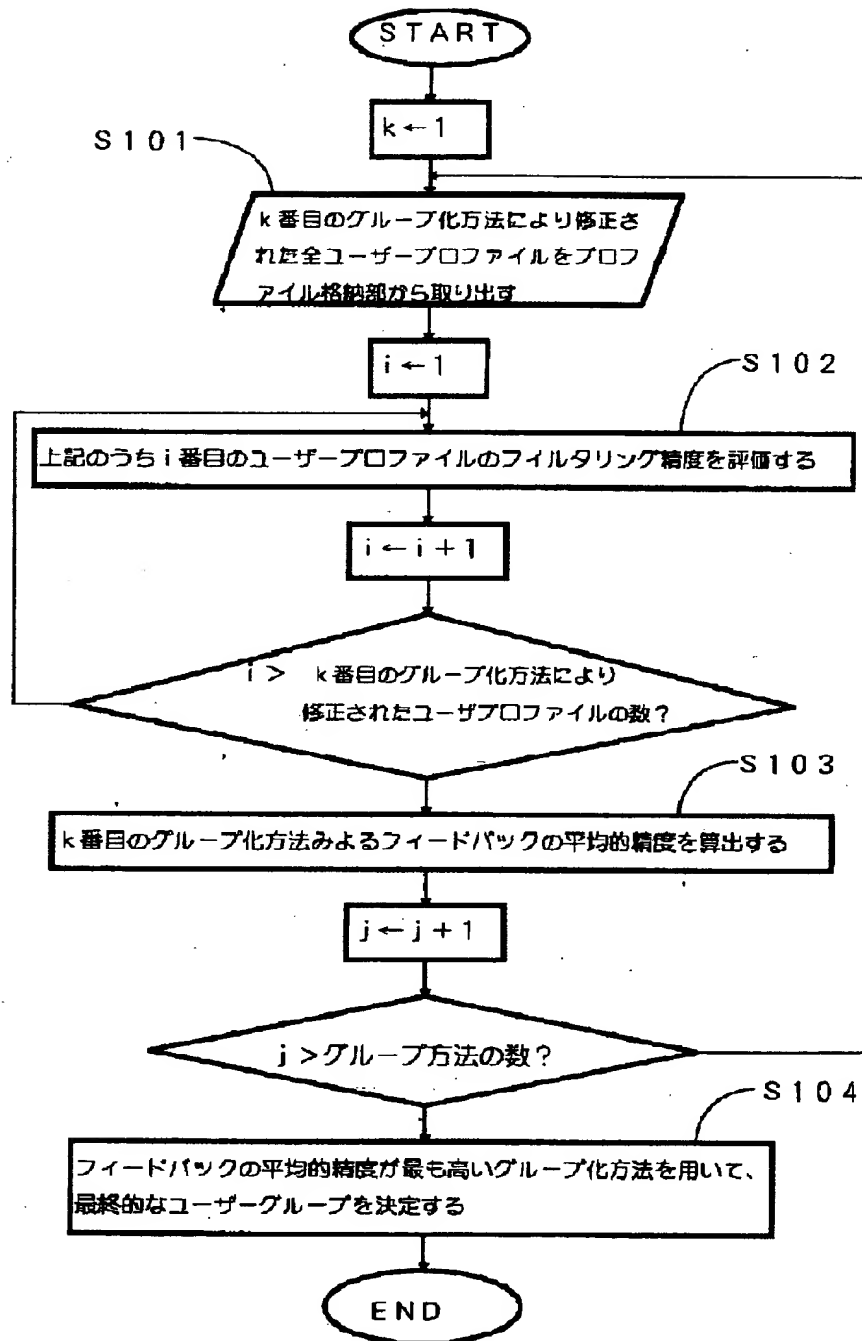
【図9】



【図11】



【図10】



フロントページの続き

(72)発明者 住田 一男  
神奈川県川崎市幸区小向東芝町1番地 株  
式会社東芝研究開発センター内

Fターム(参考) 5B075 NR10 NR12 PR03 PR06 PR08  
UU40